Mathematical models for the geographic profiling problem

Mike O’Leary

Department of Mathematics
Towson University

Georgetown University
Mathematics Department Colloquium
March 17, 2009
Acknowledgements

- Towson University Applied Mathematics Laboratory
  - Undergraduate research projects in mathematics sponsored by companies and government agencies
  - Student participants in 2005-06, 2006-07, and 2007-08:
    - Lauren Amrhine, Brooke Belcher, Brandie Biddy, Colleen Carrion, Chris Castillo, Paul Corbitt, Gregory Emerson, Adam Fojtik, Yu Fu, Natasha Gikunju, Laurel Mount, Kristopher Seets, Jonathan Vanderkolk, Grant Warble, Ruozhen Yao, Melissa Zimmerman
  - Faculty participants:
    - Coy L. May (Towson University) (2005-2007)
    - Andrew Engel (SAS) (2005)

- National Institute of Justice
  - Supported by 2005-IJ-CX-K036 and 2007-DE-BX-K005
  - Special thanks to Ron Wilson for his support
  - Thanks to Stanley Erickson, Iara Infosino, and Tommy Sexton

- Phil Canter, Baltimore County Police Department
The Geographic Profiling Problem

FIELDS ARRANGED BY PURITY

SOCIOLGY IS JUST APPLIED PSYCHOLOGY

PSYCHOLOGY IS JUST APPLIED BIOLOGY.

BIOLOGY IS JUST APPLIED CHEMISTRY

WHICH IS JUST APPLIED PHYSICS. IT'S NICE TO BE ON TOP.

OH, HEY, I DIDN'T SEE YOU GUYS ALL THE WAY OVER THERE.

SOCIOLGISTS  PSYCHOLOGISTS  BIOLOGISTS  CHEMISTS  PHYSICISTS  MATHEMATICIANS
The Geographic Profiling Problem

How can we estimate for the location of the anchor point of a serial offender from knowledge of the locations of the offender’s crime sites?

- The anchor point can be the offender’s place of residence, place of work, or some other location important to the offender.
<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Location</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>March 8</td>
<td>12:30 pm</td>
<td>-76.71350 39.29850</td>
<td>Speedy Mart</td>
</tr>
<tr>
<td>March 19</td>
<td>4:30 pm</td>
<td>-76.74986 39.31342</td>
<td>Exxon</td>
</tr>
<tr>
<td>March 21</td>
<td>4:00 pm</td>
<td>-76.76204 39.34100</td>
<td>Exxon</td>
</tr>
<tr>
<td>March 27</td>
<td>2:30 pm</td>
<td>-76.71350 39.29850</td>
<td>Speedy Mart</td>
</tr>
<tr>
<td>April 15</td>
<td>4:00 pm</td>
<td>-76.73719 39.31742</td>
<td>Citgo</td>
</tr>
<tr>
<td>April 28</td>
<td>5:00 pm</td>
<td>-76.71350 39.29850</td>
<td>Speedy Mart</td>
</tr>
</tbody>
</table>
What qualities should a good geographic profiling method possess?

1. The method should be logically rigorous
The Geographic Profiling Problem

What qualities should a good geographic profiling method possess?

1. The method should be logically rigorous
2. There should be explicit connections between assumptions on offender behavior and the components of the model
The Geographic Profiling Problem

What qualities should a good geographic profiling method possess?

1. The method should be logically rigorous.
2. There should be explicit connections between assumptions on offender behavior and the components of the model.
3. The method should be able to take into account local geographic features.
The Geographic Profiling Problem

What qualities should a good geographic profiling method possess?

1. The method should be logically rigorous
2. There should be explicit connections between assumptions on offender behavior and the components of the model
3. The method should be able to take into account local geographic features
   - It should be able to account for geographic features that influence the selection of a crime site
What qualities should a good geographic profiling method possess?

1. The method should be logically rigorous
2. There should be explicit connections between assumptions on offender behavior and the components of the model
3. The method should be able to take into account local geographic features
   - It should be able to account for geographic features that influence the selection of a crime site
   - It should be able to account for geographic features that influence the potential anchor points of offenders
What qualities should a good geographic profiling method possess?

1. The method should be logically rigorous.
2. There should be explicit connections between assumptions on offender behavior and the components of the model.
3. The method should be able to take into account local geographic features:
   - It should be able to account for geographic features that influence the selection of a crime site.
   - It should be able to account for geographic features that influence the potential anchor points of offenders.
4. The method should be based on data that is available to the jurisdiction(s) where the offenses occur.
What qualities should a good geographic profiling method possess?

1. The method should be logically rigorous
2. There should be explicit connections between assumptions on offender behavior and the components of the model
3. The method should be able to take into account local geographic features
   - It should be able to account for geographic features that influence the selection of a crime site
   - It should be able to account for geographic features that influence the potential anchor points of offenders
4. The method should be based on data that is available to the jurisdiction(s) where the offenses occur
5. The method should return a prioritized search area for law enforcement officers
A number of software packages have been developed to help this problem.

- CrimeStat (Ned Levine)
- Dragnet (David Canter)
- Rigel (Kim Rossmo)
A number of software packages have been developed to help this problem.

- CrimeStat (Ned Levine)
- Dragnet (David Canter)
- Rigel (Kim Rossmo)

There are a number of controversies surrounding geographic profiling.
Existing Methods

- A number of software packages have been developed to help this problem.
  - CrimeStat (Ned Levine)
  - Dragnet (David Canter)
  - Rigel (Kim Rossmo)
- There are a number of controversies surrounding geographic profiling.
  - How should you evaluate the effectiveness of a geographic profiling tool?
Existing Methods

- A number of software packages have been developed to help this problem.
  - CrimeStat (Ned Levine)
  - Dragnet (David Canter)
  - Rigel (Kim Rossmo)
- There are a number of controversies surrounding geographic profiling.
  - How should you evaluate the effectiveness of a geographic profiling tool?
Some researchers suggest that the best solution is simply to provide humans with some simple heuristics.

Existing Methods

- Some researchers suggest that the best solution is simply to provide humans with some simple heuristics.
Existing Methods

- Spatial distribution strategies

Estimate the anchor point with the centroid of the crime series locations

Estimate the anchor point with the center of minimum distance from the crime locations

Canter’s Circle hypotheses:
- The anchor point is contained in a circle whose diameter is formed by the two crime locations that are farthest apart


Probability distance strategies
- These have been implemented in software
  - CrimeStat (Ned Levine)
  - Dragnet (David Canter)
  - Rigel (Kim Rossmo)

Mike O’Leary (Towson University)  Geographic Profiling  Georgetown University
Existing Methods

- Spatial distribution strategies
  - Estimate the anchor point with the centroid of the crime series locations

Canter’s Circle hypotheses:
- The anchor point is contained in a circle whose diameter is formed by the two crime locations that are farthest apart.


Probability distance strategies
- These have been implemented in software
  - CrimeStat (Ned Levine)
  - Dragnet (David Canter)
  - Rigel (Kim Rossmo)

Mike O’Leary (Towson University)
Existing Methods

- **Spatial distribution strategies**
  - Estimate the anchor point with the centroid of the crime series locations
  - Estimate the anchor point with the center of minimum distance from the crime locations

Canter's Circle hypotheses:
- The anchor point is contained in a circle whose diameter is formed by the two crime locations that are farthest apart


- **Probability distance strategies**
  - These have been implemented in software
    - CrimeStat (Ned Levine)
    - Dragnet (David Canter)
    - Rigel (Kim Rossmo)
Existing Methods

- Spatial distribution strategies
  - Estimate the anchor point with the centroid of the crime series locations
  - Estimate the anchor point with the center of minimum distance from the crime locations
  - Canter’s Circle hypotheses:

Existing Methods

- **Spatial distribution strategies**
  - Estimate the anchor point with the centroid of the crime series locations
  - Estimate the anchor point with the center of minimum distance from the crime locations
  - **Canter’s Circle hypotheses:**
    - The anchor point is contained in a circle whose diameter is formed by the two crime locations that are farthest apart
Existing Methods

- **Spatial distribution strategies**
  - Estimate the anchor point with the centroid of the crime series locations
  - Estimate the anchor point with the center of minimum distance from the crime locations
  - Canter’s Circle hypotheses:
    - The anchor point is contained in a circle whose diameter is formed by the two crime locations that are farthest apart

- **Probability distance strategies**
Existing Methods

- **Spatial distribution strategies**
  - Estimate the anchor point with the centroid of the crime series locations
  - Estimate the anchor point with the center of minimum distance from the crime locations
  - Canter’s Circle hypotheses:
    - The anchor point is contained in a circle whose diameter is formed by the two crime locations that are farthest apart

- **Probability distance strategies**
  - These have been implemented in software
    - CrimeStat (Ned Levine)
    - Dragnet (David Canter)
    - Rigel (Kim Rossmo)
Developing a Model

To understand how we might proceed let us begin by adopting some common notation

- A point $x$ will have two components $x = (x^{(1)}, x^{(2)})$.
  - These can be latitude and longitude
  - These can be the distances from a pair of reference axes
- The series consists of $n$ crimes at the locations $x_1, x_2, \ldots, x_n$
- The offender’s anchor point will be denoted by $z$.
- Distance between the points $x$ and $y$ will be $d(x, y)$. 
How should we measure distances?

- The Euclidean distance $d_2(x, y) = \sqrt{(x^{(1)} - y^{(1)})^2 + (x^{(2)} - y^{(2)})^2}$
How should we measure distances?

- The Manhattan distance $d_1(x, y) = |x^{(1)} - y^{(1)}| + |x^{(2)} - y^{(2)}|$
How should we measure distances?

- The highway distance?
How should we measure distances?

- The street distance?
Mathematical Review of Existing Methods

- Existing algorithms begin by first making a choice of distance metric $d$; they then select a decay function $f$ and construct a hit score function $S(y)$ by computing

$$S(y) = \sum_{i=1}^{n} f(d(x_i, y)) = f(d(x_1, y)) + \cdots + f(d(x_n, y)).$$

- Essentially, a probability density function is centered at each crime site, and the result summed.

- Regions with a high hit score are considered to be more likely to contain the offender’s anchor point $z$ than regions with a low hit score.
Rossmo’s method:
- The distance metric is the Manhattan distance
- The distance decay function $f$ is

$$f(d) = \begin{cases} \frac{k}{d^h} & \text{if } d > B, \\ kB^{g-h} & \text{if } d \leq B. \end{cases}$$

Mathematical Review of Existing Methods

Canter’s method:
- The distance metric is the Euclidean distance
- The decay function is either \( f(d) = e^{-\beta d} \) or
  \[
  f(d) = \begin{cases} 
  0 & \text{if } d < A, \\
  b & \text{if } A \leq d < B, \\
  Ce^{-\beta d} & \text{if } d \geq B. 
  \end{cases}
  \]

Mathematical Review of Existing Methods

Levine’s method:

- The distance metric is the Euclidean distance
- The decay function can be linear, exponentially decaying, truncated exponentially decaying, normal, lognormal, or a function fit to decay data.

- The latest version of CrimeStat (3.1) has a new Bayesian algorithm, significantly different from this approach.

Suppose that we know nothing about the offender, only that the offender chooses to offend at the location $x$ with probability density $P(x)$.

- The probability density does not mean that the offender chooses randomly (though he may), rather we are modeling our lack of complete information about the offender.
- Probabilistic models are common in modeling deterministic phenomena, including
  - The stock market
  - Population dynamics
  - Genetics
  - Epidemiology
  - Heat flow
On what variables should the probability density $P(x)$ depend?

- The anchor point $z$ of the offender
  - Each offender needs to have a unique anchor point
  - The anchor point must have a well-defined meaning—e.g., the offender’s place of residence
  - The anchor point needs to be stable during the crime series
- The average distance $\alpha$ the offender is willing to travel from their anchor point
  - Different offender’s have different levels of mobility—an offender will need to travel farther to commit some types of crimes (e.g., liquor store robberies, bank robberies) than others (e.g., residential burglaries)
  - This varies between offenders
  - This varies between crime types
- Other variables can be included

We are left with the assumption that an offender with anchor point $z$ and mean offense distance $\alpha$ commits an offense at the location $x$ with probability density $P(x \mid z, \alpha)$
Our mathematical problem then becomes the following:

- Given a sample \( x_1, x_2, \ldots, x_n \) (the crime sites) from a probability distribution \( P(x \mid z, \alpha) \), estimate the parameter \( z \) (the anchor point).
- This is a well-studied mathematical problem.
- One approach is the theory of *maximum likelihood*.

Construct the likelihood function

\[
L(y, a) = \prod_{i=1}^{n} P(x_i \mid y, a) = P(x_1 \mid y, a) \cdots P(x_n \mid y, a)
\]

Then the best choice of \( z \) is the choice of \( y \) that makes the likelihood as large as possible.

This is equivalent to maximizing the log-likelihood

\[
\lambda(y, a) = \sum_{i=1}^{n} \ln P(x_i \mid y, a) = \ln P(x_1 \mid y, a) + \cdots + \ln P(x_n \mid y, a)
\]

The log-likelihood has a similar structure to the hit score method.
Bayesian Analysis

- Suppose that there is only one crime site $x$. Then Bayes’ Theorem implies that

$$P(z, \alpha | x) = \frac{P(x | z, \alpha) \pi(z, \alpha)}{P(x)}$$

- $P(z, \alpha | x)$ is the *posterior* distribution
  - It gives the probability density that the offender has anchor point $z$ and the average offense distance $\alpha$, given that the offender has committed a crime at $x$

- $\pi(z, \alpha)$ is the *prior* distribution.
  - It represents our knowledge of the probability density for the anchor point $z$ and the average offense distance $\alpha$ before we incorporate information about the crime.
  - If we assume that the choice of anchor point is independent of the average offense distance, we can write

$$\pi(z, \alpha) = H(z)M(\alpha)$$

where $H(z)$ is the prior distribution of anchor points, and $M(\alpha)$ is the prior distribution of mean offense distances.

- $P(x) = \int \int P(x | z, \alpha) \pi(z, \alpha) \, dz \, d\alpha$ is the *marginal* distribution
Bayesian Analysis

- A similar analysis holds when there is a series of \( n \) crimes; in this case

\[
P(z, \alpha \mid x_1, \ldots, x_n) = \frac{P(x_1, \ldots, x_n \mid z, \alpha) \pi(z, \alpha)}{P(x_1, \ldots, x_n)}.
\]

- If we assume that the offender’s choice of crime sites are mutually independent, then

\[
P(x_1, \ldots, x_n \mid z, \alpha) = P(x_1 \mid z, \alpha) \cdots P(x_n \mid z, \alpha)
\]


giving us the relationship

\[
P(z, \alpha \mid x_1, \ldots, x_n) \propto P(x_1 \mid z, \alpha) \cdots P(x_n \mid z, \alpha) H(z) M(\alpha).
\]

- Because we are only interested in the location of the anchor point, we take the conditional distribution with respect to \( \alpha \) to obtain the following
Fundamental Result

Suppose that an unknown offender has committed crimes at $x_1, x_2, \ldots, x_n$, and that

- The offender has a unique stable anchor point $z$.
- The offender chooses targets to offend according to the probability density $P(x | z, \alpha)$ where $\alpha$ is the average distance the offender is willing to travel.
- The target locations in the series are chosen independently.
- The prior distribution of anchor points is $H(z)$, the prior distribution of the mean offense distance is $M(\alpha)$ and these are independent of one another.

Then the probability density that the offender has anchor point at the location $z$ satisfies:

$$P(z | x_1, \ldots, x_n) \propto \int_0^\infty P(x_1 | z, \alpha) \cdot \ldots \cdot P(x_n | z, \alpha) H(z) M(\alpha) \, d\alpha$$
For the mathematics to be useful, we need to be able to:

- Make some reasonable choice for our model for offender behavior
- Make some reasonable choice for the prior distribution of anchor points
- Make some reasonable choice for the prior distribution of the average offense distance, and
- Be able to evaluate the mathematical terms that appear
Models of Offender Behavior

- Suppose that we assume that offenders choose offense sites according to a normal distribution, so that

\[ P(x \mid z, \alpha) = \frac{1}{4\alpha^2} \exp \left( -\frac{\pi}{4\alpha^2} |x - z|^2 \right). \]

- If we also assume that all offenders have the same average offense distance \( \tilde{\alpha} \), and that all anchor points are equally likely, then

\[ P(z \mid x_1, \ldots, x_n) = \left( \frac{1}{4\tilde{\alpha}^2} \right)^n \exp \left( -\frac{\pi}{4\tilde{\alpha}^2} \sum_{i=1}^{n} |x_i - z|^2 \right). \]

- The mode of this distribution- the point most likely to be the offender’s anchor point- is the mean center of the crime site locations.
Models of Offender Behavior

- Suppose that we assume that offenders choose offense sites according to a negative exponential distribution, so that

\[ P(x \mid z, \alpha) = \frac{2}{\pi \alpha^2} \exp \left( -\frac{2}{\alpha} |x - z| \right). \]

- If we also assume that all offenders have the same average offense distance \( \tilde{\alpha} \), and that all anchor points are equally likely, then

\[ P(z \mid x_1, \ldots, x_n) = \left( \frac{2}{\pi \tilde{\alpha}^2} \right)^n \exp \left( -\frac{2}{\tilde{\alpha}} \sum_{i=1}^{n} |x_i - z| \right) \]

- The mode of this distribution- the point most likely to be the offender’s anchor point- is the center of minimum distance of the crime site locations.
Models of Offender Behavior

- What would a more realistic model for offender behavior look like?
  - Consider a model in the form

  \[ P(x \mid z, \alpha) = D(d(x, z), \alpha) \cdot G(x) \cdot N(z) \]

- \( D \) models the effect of distance decay using the distance metric \( d(x, z) \)
  - We can specify a normal decay, so that \( D(d, \alpha) = \frac{1}{4\alpha^2} \exp \left( -\frac{\pi}{4\alpha^2} d^2 \right) \)
  - We can specify a negative exponential decay, so that \( D(d, \alpha) = \frac{2}{\pi\alpha^2} \exp \left( -\frac{2}{\alpha} d \right) \)
  - Any choice can be made for the distance metric (Euclidean, Manhattan, et.al)

- \( G \) models the geographic features that influence crime site selection
  - High values for \( G(x) \) indicate that \( x \) is a likely target for typical offenders;
  - Low values for \( G(x) \) indicate that \( x \) is a less likely target

- \( N \) is a normalization factor, required to ensure that \( P \) is a probability distribution
  - \( N(z) = \left[ \int \int D(d(y, z), \alpha) G(y) \, dy^{(1)} \, dy^{(2)} \right]^{-1} \)
  - \( N \) is completely determined by the choices for \( D \) and \( G \).
Geographic Features that Influence Crime Selection

- G models the geographic features that influence crime site selection, with high values indicating the location was more likely to be targeted by an offender.
- How can we calculate G?
  - Use available geographic and demographic data and the correlations between crime rates and these variables that have already been published to construct an appropriate choice for $G(x)$
    - Different crime types have different etiologies; in particular their relationship to the local geographic and demographic backcloth depends strongly on the particular type of crime. This would limit the method to only those crimes where this relationship has been well studied
  - Some crimes can only occur at certain, well-known locations, which are known to law enforcement
    - For example, gas station robberies, ATM robberies, bank robberies, liquor store robberies
    - This does not apply to all crime types- e.g. street robberies, vehicle thefts.
  - We can assume that historical crime patterns are good predictors of the likelihood that a particular location will be the site of a crime.
Convenience Store Robberies, Baltimore County

[Map showing the distribution of convenience store robberies in Baltimore County]
Suppose that historical crimes have occurred at the locations $c_1, c_2, \ldots, c_N$.

Choose a kernel density function $K(y | \lambda)$

- $\lambda$ is the bandwidth of the kernel density function

Calculate $G(x) = \sum_{i=1}^{N} K(d(x, c_i) | \lambda)$

- The bandwidth $\lambda$ can be e.g. the mean nearest neighbor distance
- Effectively this places a copy of the kernel density function on each crime site and sums
Convenience Store Robberies, Baltimore County
Distance Decay: Buffer Zones

- A buffer zone is a region around the offender’s anchor point where they are less likely to offend, presumably due to a fear of being recognized.

- Consider the following models of offender behavior:

- Which shows evidence of a buffer zone?
A buffer zone is a region around the offender’s anchor point where they are less likely to offend, presumably due to a fear of being recognized.

Consider the following models of offender behavior:

- Which shows evidence of a buffer zone?
  - These are two views of the same distribution
Distance Decay: Buffer Zones

- A buffer zone is a region around the offender’s anchor point where they are less likely to offend, presumably due to a fear of being recognized.

- Consider the following models of offender behavior:

  - Which shows evidence of a buffer zone?
    - These are two views of the same distribution
    - If the offender is using a two-dimensional normal distribution to select targets, then the appropriate distribution for the offense distance is the Rayleigh distribution.
Distance Decay

- Suppose that the (two-dimensional) distance decay component $D(d(x, z) | \alpha)$ is modeled with a Euclidean distance $d$
- Then the (one-dimensional) distribution of offense distances $D_{\text{one-dim}}(d | \alpha)$ is given by

$$D_{\text{one-dim}}(d | \alpha) = 2\pi d \cdot D(d | \alpha)$$

- In particular, $D_{\text{one-dim}}(d | \alpha) \to 0$ as $d \to 0$, regardless of the particular choice of $D(d | \alpha)$, provided $D(0 | \alpha) < \infty$. 

When considering the effect of distance, it is essential to be aware of the dimension of the underlying function.
Distance Decay

- Suppose that the (two-dimensional) distance decay component $D(d(x, z) \mid \alpha)$ is modeled with a Euclidean distance $d$.

- Then the (one-dimensional) distribution of offense distances $D_{\text{one-dim}}(d \mid \alpha)$ is given by

$$D_{\text{one-dim}}(d \mid \alpha) = 2\pi d \cdot D(d \mid \alpha)$$

- In particular, $D_{\text{one-dim}}(d \mid \alpha) \to 0$ as $d \to 0$, regardless of the particular choice of $D(d \mid \alpha)$, provided $D(0 \mid \alpha) < \infty$.

- When considering the effect of distance, it is essential to be aware of the dimension of the underlying function.
Distance Decay: Residential Burglaries in Baltimore County

Baltimore County Residential Burglaries 1991-2008
Distance Decay: Non-Residential Burglaries in Baltimore County
Distance Decay: Data Fitting

- Suppose that we measure the aggregate number of offenders who commit a crime at a distance $d$ from their anchor point; call the relative fraction $A(d)$.

- Different offenders are willing to travel different distances to offend; $M(\alpha)$ was defined to be the probability distribution for the mean offense distance across offenders.

- Suppose that each offender chooses targets according to $D_{\text{one-dim}}(d \mid \alpha)$

- Then
  
  $$A(d) = \int_0^\infty D_{\text{one-dim}}(d \mid \alpha)M(\alpha)\,d\alpha$$

- Since $A(d)$ can be measured and $D_{\text{one-dim}}(d \mid \alpha)$ modeled, we can solve this equation for the prior mean offense travel distance $M(\alpha)$
The operator $M \mapsto A$ given by

$$A(d) = \int_0^\infty D_{\text{one-dim}}(d \mid \alpha) M(\alpha) \, d\alpha$$

is smoothing; we expect that the inverse operator $A \mapsto M$ to be ill-posed.

If we choose a normal form for the two-dimensional decay function (and so a Rayleigh form for the one-dimensional decay function), then

$$A(d) = \int_0^\infty \frac{\pi d}{2\alpha^2} \exp \left(-\frac{\pi d^2}{4\alpha^2}\right) M(\alpha) \, d\alpha$$

If we make the changes of variables $p = \pi/4\alpha^2$, $\alpha = \sqrt{\pi/2p}$, $\omega(p) = \alpha M(\alpha)$, $s = d^2$, we obtain

$$\frac{A(\sqrt{s})}{\sqrt{s}} = \int_0^\infty e^{-sp} \omega(p) \, dp = \mathcal{L}(\omega)(s)$$
Choose a step size $\delta > 0$, and suppose choose $N$ so that
- $A(d) \approx 0$ for $d \geq N\delta$; then
- $M(d) \approx 0$ for $d \geq N\delta$.

Suppose that $A(d)$ is not known exactly, but that a sample
\{$\rho_1, \rho_2 \ldots, \rho_S$\} of size $S$ has been drawn.
- Define $a_j = \# \{s \mid d_{j-1} \leq \rho_s < d_j\}$
- Then $A(d_j)\delta \approx a_j/S$

Apply collocation at the points $d^*_k = (k + \frac{1}{2})\delta$, $1 \leq k \leq N$ and approximate the integral by the midpoint rule at the nodes
\[ \alpha^*_j = (j + \frac{1}{2})\delta, \quad 1 \leq j \leq N, \]
to find the linear discretization of the integral equation
\[ a = Gm \]

- $G = G_{jk} = \frac{\pi S\delta}{2} \frac{(j - \frac{1}{2})}{(k - \frac{1}{2})^2} \exp \left( -\frac{\pi}{4} \frac{(j - \frac{1}{2})^2}{(k - \frac{1}{2})^2} \right)$
- $a = (a_1, a_2, \ldots, a_N)$
- $m = (M(\alpha^*_1), M(\alpha^*_2), \ldots, M(\alpha^*_N))$
Attempts to directly solve the equation $Gm = a$ for $m$ fail due to numerical instability; though $G$ is analytically non-singular, it is not numerically non-singular.

Attempts to solve the equation using the pseudo-inverse $G^\dagger$ so that $m = G^\dagger a$ still fail due to numerical instability.

Write $G = USV^\top$ with $S = \text{diag}(s_1, s_2, \ldots, s_N)$, then $s_j \to 0$ with no appreciable gaps.

$G$ has ill-defined numerical rank.

We can apply Tikhonov regularization; i.e. replace $S^\dagger$ with

$$S^\dagger_\lambda = \text{diag}\left(\frac{s_1}{s_1^2 + \lambda^2}, \frac{s_2}{s_2^2 + \lambda^2}, \ldots, \frac{s_N}{s_N^2 + \lambda^2}\right)$$

then $m = G^\dagger_\lambda a$ can be calculated.
Distance Decay: Solving the Integral Equation

OH, YOU'RE NOT FOOLING ANYBODY! YOU'RE JUST MAKING STUFF UP NOW!

\[ \frac{2x}{x^2} \frac{x-x_0}{x-0} \]

\[ \frac{x^5}{x^2} = \frac{3}{y} \]

\[ \frac{y}{x} = \frac{2 + 2x}{x} \]

\[ \frac{y}{x} = \frac{2 + 2x}{x} \]
Distance Decay: Residential Burglaries in Baltimore County - Model Fit
Distance Decay: Non-Residential Burglaries in Baltimore County - Model Fit
Distance Decay: Residential Burglaries in Baltimore County- Prior Distribution

Average distance for residential burglary in Baltimore County
Distance Decay: Residential Burglaries in Baltimore County- Prior Distribution by Age
Anchor Points

- We have assumed
  - Each offender has a unique, well-defined anchor point that is stable throughout the crime series
  - The function $H(z)$ represents our prior knowledge of the distribution of anchor points before we incorporate information about the crime series.

- What are reasonable choices for the anchor point?
  - Residences
  - Places of work

- Suppose that anchor points are residences - can we estimate $H(z)$?
  - Population density information is available from the U.S. Census at the block level, sorted by age, sex, and race/ethnic group.
    - We can use available demographic information about the offender
    - Set $H(z) = \sum_{i=1}^{N_{\text{blocks}}} p_i K(z - q_i | \sqrt{A_i})$
    - Here block $i$ has population $p_i$, center $q_i$, and area $A_i$.

- Distribution of residences of past offenders can be used.
  - Calculate $H(z)$ using the same techniques used to calculate $G(x)$
Washington D.C., 18-29 year old white non-hispanic men
Washington D.C., 18-29 year old white hispanic men
Washington D.C., 18-29 year old black men
Code that implements this method is nearing completion.

Required Input:
- Crime series locations
- Representative selection of the locations of historically similar crimes, (as determined by the analyst) to estimate target attractiveness
- Geographic boundaries of the jurisdiction(s) reporting the crime series and historical crimes
- Available demographic information about the offender, if any
- Locations of both anchor points and crime sites of historically similar crimes (as determined by the analyst) to estimate the distribution of average offense distances

The code will then automatically
- Calculate an estimate of the target attractiveness distribution
- Estimate the prior distribution of anchor points, assuming anchor point density is proportional to population density
- Estimate the prior distribution of average offense distances

The code will then return a map giving the probability distribution for offender anchor points
- Available output formats include .kml and .csv, for display and analysis is a wide range of further applications.
Questions?

Mike O’Leary
Department of Mathematics
Towson University

moleary@towson.edu